ABSOLUTE ENDPOINT DATA DISCOVERY

FREQUENTLY ASKED QUESTIONS

OVERVIEW

What is Endpoint Data Discovery?

Endpoint Data Discovery is a unique data security feature available within the Absolute platform. Endpoint Data Discovery is a key feature of Absolute's Data Visibility and Protection solution. This capability allows you to set policies to scan your managed Windows and Mac devices for data-at-risk.

How does Endpoint Data Discovery work?

This feature is policy-based and can be configured within Device Policies in the Absolute platform. When the Endpoint Data Discovery policy is configured and activated, the files stored on your managed devices are scanned to identify at-risk data such as:

- Credit card numbers
- Social security numbers
- Personal health information, such as a patient's name, date of birth, or medical diagnosis
- Personal financial information, such as a bank account number
- Encrypted or password-protected files
- Custom information that is unique to your organization

Two types of scans can be set and run at scheduled frequencies. Full scans scan all content in defined locations, while Delta scans scan files that have been added or modified since the last full scan. Device policies allow scans to target specific devices, or be applied more broadly. Scan results are presented in customizable reports within the Absolute console. You can review these reports to identify at-risk devices, and then initiate actions to remediate where necessary.

How do I activate Endpoint Data Discovery in the Absolute platform?

Simply configure Endpoint Data Discovery as a Device Policy.

What happens once Endpoint Data Discovery is activated?

The Absolute agent is responsible for silently scanning files stored on your devices to detect content that is confidential or at risk. When you activate an Endpoint Data Discovery policy, a small component is deployed automatically to each device after the next successful agent call. A full scan of each device is performed. During the scan, the Endpoint Data Discovery component opens each file on the hard drive, scans files for specific pieces of information (matched tokens), masks tokens where applicable, encrypts these matched tokens, and uploads them to the Absolute Monitoring Center using a secure connection.

What Absolute product editions receive Endpoint Data Discovery?

Endpoint Data Discovery is available with the Data Awareness functionality in the Absolute Platform. Absolute Resilience customers recieve Data Awareness, and it is also available to Absolute Visibility and Absolute Control customers as a product add-on module. Absolute DDS Premium, Absolute DDS for Education Premium and Absolute DDS for Healthcare Premium. Legacy Absolute DDS Premium, Computrace Complete and Computrace One licenses also receive Endpoint Data Discovery.

How can I evaluate the Absolute platform and Data Awareness?

Contact an Absolute Sales Representative to organize an evaluation.

DEVICE POLICIES

How do I apply Endpoint Data Discovery in the Absolute console?
By default, the global policy group includes a preconfigured Endpoint Data Discovery policy, which is set to inactive. In all likelihood, you will want to apply different policies to subsets of your devices. Therefore, it's best practice to create multiple customized policy groups and define a unique policy for each group of devices.

How do I configure an Endpoint Data Discovery policy?
To configure an Endpoint Data Discovery policy you need to complete three main tasks:

1. **Set the Endpoint Data Discovery rules:** Select the confidential or at-risk data patterns to search for
2. **Set the scan schedule:** Regular or one-off scans with the ability to schedule scans for a specific time and date
3. **Set the scan level:** The scope of the scan, targeted to a subset of file types or a comprehensive scan across multiple file types

ENDPOINT DATA DISCOVERY SCANS

How frequently does Absolute scan devices?
The scan frequency is fully customizable. Two different scan types are available, a full scan and a delta scan.

SCAN TYPE	DESCRIPTION
Full Scan	Opens and scans the files on a hard drive according to the scan configurations. Full scans can occur on a monthly or quarterly basis.
Delta Scan	Opens and scans the files on a hard drive that were added or edited since the last full scan. Delta scans can occur on a daily or a weekly basis.

Best practice is to set a schedule that uses a combination of full and delta scans. For example, you may want to use one of the following scan schedules:

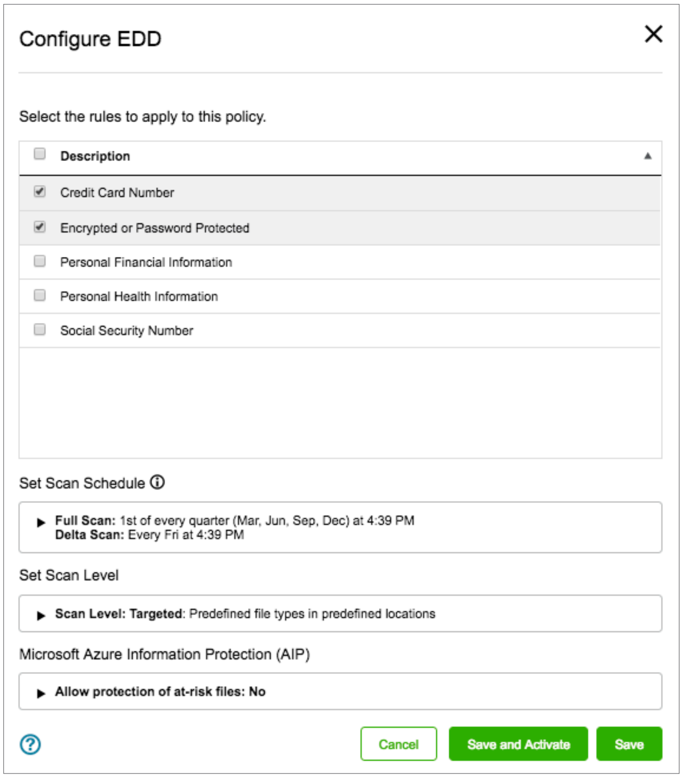
- Monthly full scans and twice weekly delta scans
- Quarterly full scans and weekly delta scans

When can scans be scheduled?
You can determine the time and day full or delta scans occur.

How much historical scan data is retained?
Absolute retains data from the last two full scans and all subsequent delta scans. All prior data is discarded.

Can the scan identify content inside zip files?
Yes, Absolute Endpoint Data Discovery can scan and identify at-risk data within zip files.

How do you scan encrypted files?
Absolute identifies encrypted files and password-protected files. However because they are encrypted or password-protected, they cannot be opened. Absolute does, however, assign the Encrypted content type to these file types to advise on the number of encrypted files on an endpoint. In reporting, each encrypted or password-protected file generates a Match Score of 1.



Configure EDD

Select the rules to apply to this policy.

□ **Description**

☑ Credit Card Number

☑ Encrypted or Password Protected

□ Personal Financial Information

□ Personal Health Information

□ Social Security Number

Set Scan Schedule

Full Scan: 1st of every quarter (Mar, Jun, Sep, Dec) at 4:39 PM
Delta Scan: Every Fri at 4:39 PM

Set Scan Level

Scan Level: Targeted: Predefined file types in predefined locations

Microsoft Azure Information Protection (AIP)

Allow protection of at-risk files: No

Figure 1: Configuring Data Scans

What is the performance impact of scans?
Absolute has tested Endpoint Data Discovery scans extensively to ensure minimum device performance impact. This feature was used on 10,000's of devices during the beta program participants did not report any noticeable negative performance impact. Scans are run with background priority in Windows, which means they are relatively inactive when devices are in use, and take advantage of computing cycles when available to perform scans.

How long do scans take to run?
This varies depending on the device and amount of data. Full scans can take between a few hours and a few days, depending on the number of devices, the number of files on each hard drive, and the size of those files. Delta scans are much shorter and only scan files that have been modified or created since the last full scan.

Do scans require user interaction / does a user know a scan is running?
No, scans were designed to require no user interaction. This is beneficial when investigating potential insider threats.

What file types are scanned?
Archive files, database files, email and messaging formats, office formats, presentation formats, spreadsheet formats, text and markup formats, word processing formats and more. View the full list in the **Appendix**.

Can I specify different scan levels?
Yes, two scan level types are available so you can define which file types and file locations to scan for content that is confidential or at risk. You can select one of the following levels:

SCAN LEVEL	DESCRIPTION
Targeted (default)	This level is associated with the narrowest scan scope. When this option is selected, files stored on a device in directories typically used to store user data are scanned. Directories that are typically used to store program files and system files are not scanned. This level is also limited to file types that are usually used to store data, see **Appendix 'Targeted Scan'** Due to the narrow scope, targeted scans tend to take less time to complete than the other options. They may also generate fewer false positives.
Moderate	This level is associated with a broader scan scope than a targeted scan. When this option is selected, the complete list of file types (see **Appendix**) are scanned for in directories typically used to store user data. Directories that are typically used to store program files and system files are not scanned.

ENDPOINT DATA DISCOVERY RULES

What is a custom Endpoint Data Discovery rule?

A custom Endpoint Data Discovery rule specifies terms you wish to discover in data residing on endpoints. A number of templates are offered to allow you to easily create these rules. Templates range from Social Insurance Numbers, General ID samples, health and financial terms, to help define custom terms unique to your organization. Rules are provided as expression sets, and each expression set can have multiple expressions (also known as lexicons). Custom rules can be configured under **Data Protection › Endpoint Data Discovery › Rules**

What types of Endpoint Data Discovery rule templates are included?

Five rules are provided as expression set templates to help you get started. You can build on these to create your customized rules. Each template contains expressions to detect a particular type of content:

TEMPLATE	DESCRIPTION
US Social Security Numbers	Includes expressions to detect valid United States Social Security Numbers
Canadian Social Insurance Numbers	Includes expressions to detect valid Canadian Social Insurance Numbers
General ID Samples	Includes sample expressions to detect different types of identification numbers, such as student IDs, account numbers, loan numbers, patient IDs, and more. If you use this expression set template in your rule, you will most likely need to remove the expressions that are not relevant and edit the remaining expressions to best suit your needs.
Health Terms	Includes a list of over 5400 health conditions, diagnoses, and medications (also known as a lexicon) and is synonymous with the lexicon used by the predefined Personal Health Information rule to detect at-risk file content. You can edit this expression set template to meet the specific needs of your organization.
Financial Terms	Includes a list of 130 terms related to financial information (also known as a lexicon), which is used by the predefined Personal Financial Information rule to detect at-risk file content. You can edit this expression set template to meet the specific needs of your organization.

Can specific terms or intellectual property be discovered?

Yes, this is achieved through a custom rule. These rules are created as expressions. An expression may be a single word or phrase, such as 'account number', or it may include a combination of words, variables, operators, and special characters. This is useful for discovering intellectual property, such as project names. When you add an expression to an expression set, you need to use the correct syntax. **Absolute Professional Services** is available to assist with custom rules.

REPORTING

How are the results of scans presented?
During a scan, the Absolute agent detects all file content that matches a policy rule. You can view the details about these matches in the following reports:

Endpoint Data Discovery History Report
The Endpoint Data Discovery History report provides a history of the content matches detected on your **Windows** and **Mac** devices by the policy. The report is generated as a result of the active policies within your account. The scope of the report is limited to information collected during the last **two** full scans of each device and all subsequent delta scans. If a scan is currently in progress when you run the Endpoint Data Discovery History report, any data collected up to that point is available in the report.

Endpoint Data Discovery Match Score Summary Report
The Endpoint Data Discovery Match Score Summary report provides a summary of the information collected by the policy on your **Windows** and **Mac** devices. The report is generated as a result of the active policies within your account. For example, you may have configured the rules in a policy to detect credit card information stored on a set of devices, and another policy to detect Personal Health Information found on another set of devices. A Match Score is provided for each active rule. Match Score is a computed value indicating the number of content matches detected on a device for the associated rule.

Endpoint Data Discovery Reporting Data Report
The Endpoint Data Discovery Reporting Data report presents an unfiltered view of all available information collected by the active policies within your account. By adding filters to this report, and adjusting the report's columns, you can create customized reports that include the information that is most relevant to your business needs.

You can also view this information on the Endpoint Data Discovery Summary and Endpoint Data Discovery History pages for each device.

Endpoint Data Discovery report information includes the details about the content that generated the match, the name and location of the file, and a Match Score. You can use these details to identify the devices that may require remedial action to address the at-risk data.

Can reports be customized using data from Endpoint Data Discovery scans?
Yes, using the flexible reporting in the Absolute platform, you can select information that is important to you and build and filter reports. These can also be exported for further analysis.

SECURITY

Are documents containing matched tokens stored by Absolute?
No, the documents that have been identified as containing matched tokens such as credit card numbers and social security numbers are not stored by Absolute after a match is found.

Are the matched tokens identified encrypted?
It's important to note that when Endpoint Data Discovery identifies documents containing matched tokens, those documents are not stored by Absolute.

Once matches are identified on the endpoint, the matched tokens are first encrypted at the endpoint before being sent to the Absolute console. All matched tokens are encrypted using RSA 2048 bit encryption, including those tokens matched using the built-in rules and custom rules. Additionally, any matches of credit card numbers and social security numbers using the built-in rules (Credit Card Numbers, Social Security Numbers, Personal Health Information, Personal Financial Information) are also first masked at the endpoint before being encrypted, such that only the first four digits of a match are sent and the remaining digits masked using the * symbol. Note that masking is applied to the built-in rules listed above, but not to tokens matched by custom rules. However, all matches are encrypted, regardless of whether they matched using built-in or custom rules.

Only roles of 'Security Administrator' are able to view the results of matched tokens in the Absolute console. When matches are viewed, they are only decrypted for viewing within the current session of the browser and never decrypted at the Absolute Monitoring Center.

What is the encryption level of the matched tokens?
All matched tokens are encrypted at the endpoint using RSA 2048 bit encryption. Matched tokens are only decrypted for viewing within the session of the browser, never at the Absolute Monitoring Center, and can only be viewed by the role of 'Security Administrator' within the Absolute console.

ADAPTIVE ENDPOINT SECURITY

How can Endpoint Data Discovery be used in the context of a potential security incident?
Based on the at-risk data identified on a device, organizations can take action such as remotely freezing the device or deleting the at-risk data from the device, even if the device is off the network, by way of a persistent connection that Absolute provides to the endpoint.

If at-risk data is discovered, what actions are available within the Absolute platform?
Absolute uniquely combines self-healing endpoint security with continous data visibility and protection. If at-risk data is discovered, you can initiate a Device Freeze or Data Delete with a single click. These Risk Response features allow you to protect data, or take remediation steps to prevent potential security incidents. To learn more visit absolute.com/platform

APPENDIX

TARGETED SCAN

application/CDFV2-encrypted
application/msword
application/pdf
application/vnd.ms-excel
application/vnd.ms-office
application/vnd.ms-powerpoint
application/vnd.oasis.opendocument.presentation
application/vnd.oasis.opendocument.spreadsheet
application/vnd.oasis.opendocument.text
application/x-7z-compressed
application/x-gzip
application/x-rar
application/x-tar
application/zip
text/html
text/plain
text/rtf

APPENDIX

FULL SCAN

FILE TYPE	APPLICATION OR FORMAT	EXTENSION
Archive formats	7-Zip	.7Z
	Apple Disk Image	.DMG
	ARJ	.ARJ
	Bzip2	.BZ2, .TBZ2
	ISO Disk Image	.ISO
	Java Archive	.JAR
	Microsoft Cabinet	.CAB
	Microsoft Office Binder	.OBD
	Red-Hat Package Manager	.RPM
	Roshal Archive	.RAR
	Roshal Archive (Multi-part)	
	Self-extracting.exe	.EXE
	GNU Zip	.GZ
	UNIX cpio	.CPIO
	UNIX Tar	.TAR
	Zip	.ZIP
Database formats	dBase file	.DBF
	dBASE III file	.DB, .DB3
	Paradox Database file	.DB
E-mail and messaging formats	Apple Mail	.EMLX
	Encoded mail message	.MHT
	Eudora	.MBX
	Microsoft Outlook	.MSG, .OST, .PST
	Microsoft Outlook Express	.EML
	Microsoft Outlook Forms Template	.OFT
	Sendmail "mbox"	.MBOX
	Thunderbird	
Other formats	Log File	.LOG
	Microsoft Project	.MPP, .MPX (98 only)
	Open Access II (OAII)	Not applicable*
	Uniplex	Not applicable*
	vCard	.VCF

FILE TYPE	APPLICATION OR FORMAT	EXTENSION
Presentation formats	IBM Lotus Symphony Presentation	.ODP, .SXI
	LibreOffice Presentation	.ODS
	Microsoft PowerPoint (for Windows and Mac)	.PPT, .PTTX
	OpenOffice Impress	.ODP
	StarOffice Impress 3	.SDI, .SDP, .SXI
Spreadsheet formats	Comma Separated Values	.CSV
	Framework Spreadsheet	.FW3
	IBM Lotus Symphony Spreadsheet	.ODS, .SX, .SXS
	LibreOffice Spreadsheet	.ODS
	Lotus 1-2-3	.WK, .WK3, .WK4, .WKS
	Microsoft Excel for Windows	.XLS, .XLSX, .XLSB
	Microsoft Excel for Mac	.XLS, .XLSX
	Microsoft Work SS for DOS	.WPS
	Microsoft Works SS for Windows	.ODS
	OpenOffice Calc3	
	StarOffice Calc	.ODS, .SXC, .SXS
Text and markup formats	ASCII Text	.TXT
	ANSI Text	
	HTML (Text only)	.HTM, .HTML
	HTML (Metadata Only)	
	HTML (Codes Revealed)	
	IBM DCA	.DCA, .RFT, .TXT
	Microsoft HTML Help	.CHM
	Microsoft OneNote	.ONE
	Microsoft OneNote TOC	.ONETOC
	Rich Text Format	.RTF
	SGML Text	.SGML
	Source	Not applicable*
	Transcript	
	Unicode UCS2 (big e and little e)	
	Unicode UTF8	
	Unicode UTF16 (big e and little e)	
	Windows Enhanced Meta File 1	.EMF
	Windows Meta File 1	.WMF
	XML	.XML

FILE TYPE	APPLICATION OR FORMAT	EXTENSION
Word processing and general office formats	Adobe PDF	.PDF
	Ami Pro for Windows	.AMI, .SAM
	Apple iWork	.KEY, .NUMBERS, .PAGES
	Framework WP	.FW3
	Hangul	HWP
	IBM DCA/FFT	.FFT, .RFT
	IBM DisplayWrite	.DCA, .DOC,. DW4, .DW5, .RFT
	IBM Lotus Symphony Document	.ODT
	JustSystems Ichitaro	.JBW, .JTD, .JTT
	LibreOffice Document	.ODT
	Lotus Manuscript	.MAN, .MANU, .MNU
	Mass 11	.M11
	Microsoft Publisher	.PUB
	Microsoft Word for DOS	.DOC
	Microsoft Word (for Windows or Mac)	.DOC, .DOCX
	Multimate	.DOX
	Multimate Advantage	
	OpenOffice Writer	.ODT
	Professional Write for DOS	.PW, .PWI, PW2
	Professional Write Plus for Windows	.PW
	Q&A Write	.QA, .QA3
	StarOffice Writer	.SDW, .SXW
	Wang IWP	.IWP
	Wang WP Plus	
	Windows Write	.WRI
	WinWord	.DOC
	WordPerfect (for DOS, Mac, or Windows)	.WPD
	Wordstar 2000 for DOS	.DOC, .WS2
	Wordstar for DOS	.WS, .WSX
	Wordstar for Windows	.WSD
	XYwrite	.XY

NOTE: "Not applicable*" indicates that no particular extension is associated with this application or format.